2012/06/30
Head Id: 133456
Archive Id: 133535:133915
Archive Date: 2012/06/28
Archive Tag: trunk

Measurement of the top polarization in the dilepton final state

The CMS Collaboration

Abstract

A measurement of the top quark polarization in $t\bar{t} \to \ell^+\ell^-$ events is performed in a data sample corresponding to a total integrated luminosity of $5.0\,\text{fb}^{-1}$ collected by the CMS experiment in pp collisions at a centre-of-mass energy of 7 TeV at the LHC. The measured value in the helicity basis is $P_n = -0.009 \pm 0.029 \pm 0.041$, in agreement with the standard model expectation.

1 Introduction

In the standard model (SM), tops and anti-tops in $t\bar{t}$ pair production are produced unpolarized from QCD. A small net polarization is expected in the SM from electroweak corrections to $t\bar{t}$ production. In physics beyond SM, couplings of the top to new particles can alter its polarization. Thus net polarization of the tops produced in $t\bar{t}$ production would be a good way to separate the new physics from the SM. In paper [1], the top polarization has been identified as an observable capable of distinguishing between different models which could be responsible for the large deviation of the $t\bar{t}$ forward-backward asymmentry observed in the Tevatron [2, 3].

The polarization of the top is reflected in the kinematic distributions of its daughters, because the top decays before hadronization effects can wash away this information. Among all the particles coming from the decay of the top, the charged lepton is most sensitive to the top's polarization. The top polarization P_n along a chosen axis $\hat{n}$ can thus be measured from the angular distribution of the charged leptons from the decays, measured in the top quark rest frame:

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta_{l,n}} = \frac{1}{2}(1 + 2\kappa_l P_n \cos\theta_{l,n}), \tag{1}$$

where $\theta_{l,n}$ is the direction of the lepton with respect to $\hat{n}$ and κ_l is the spin analyzing power of the lepton, equal to 1.0. In this analysis the helicity axis is used, where $\hat{n}$ is given by the direction of the top in the $t\bar{t}$ CM frame.

This note presents a measurement of the top polarization in the helicity basis, using a data sample corresponding to an integrated luminosity of 5.0 fb^{-1} collected at $\sqrt{s} = 7$ TeV by the Compact Muon Solenoid (CMS) experiment at the LHC. A detailed description of the CMS detector can be found elsewhere [4]. Dilepton decays of the $t\bar{t}$ pair are used, and the top polarization is measured using the reconstructed objects. Because the reconstructed polarization is shaped by the reconstruction efficiency and resolution, we apply an unfolding technique to recover the parton-level distribution which can be compared with theoretical predictions.

2 Event samples, reconstruction, and selection

The data used for this measurement were collected using one of the ee, eμ, or $\mu\mu$ high-p_{T} double-lepton triggers. Muon candidates are reconstructed using two algorithms that require consistent signals in the tracker and muon systems: one matches the extrapolated trajectories from the silicon tracker to signals in the muon system (tracker-based muons), and the second performs a global fit requiring consistent patterns in the tracker and the muon system (globally fitted muons) [5]. Electron candidates are reconstructed starting from a cluster of energy deposits in the electromagnetic calorimeter. The cluster is then matched to signals in the silicon tracker. A selection using electron identification variables based on shower shape and track-cluster matching is applied to the reconstructed candidates [6]. Electron candidates within $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.1$ from a muon are rejected to remove candidates due to muon bremsstrahlung and final-state radiation. Both electrons and muons are required to be isolated from other activity in the event. This is achieved by imposing a maximum allowed value of 0.15 on the ratio of the scalar sum of track transverse momenta and calorimeter transverse energy deposits within a cone of $\Delta R < 0.3$ around the lepton candidate direction at the origin (the transverse momentum of the candidate is excluded), to the transverse momentum of the candidate.

Event selection is applied to reject events other than those from $\mathrm{t\bar{t}}$ in the dilepton final state. Events are required to have two opposite-sign, isolated leptons ($\mathrm{e^+e^-}$, $\mathrm{e^\pm\mu^\mp}$, or $\mu^+\mu^-$). Both leptons must have transverse momentum $p_\mathrm{T} > 20\,\mathrm{GeV}/c$, and the electrons (muons) must have $|\eta| < 2.5$ (2.4). The reconstructed lepton trajectories must be consistent with a common interaction vertex. In the rare case ($< 0.1\%$) of events with more than two such leptons, the two leptons with the highest p_T are selected. Events with an $\mathrm{e^+e^-}$ or $\mu^+\mu^-$ pair with invariant mass between 76 and 106 GeV/c^2 or below 12 GeV/c^2 are removed to suppress Drell–Yan (DY) events ($\mathrm{Z}/\gamma^* \to \ell^+\ell^-$) as well as low mass dilepton resonances. The jets and the missing transverse energy $E_\mathrm{T}^\mathrm{miss}$ are reconstructed with a particle-flow technique [7]. The anti-k_T clustering algorithm [8] with a distance parameter of 0.5 is used for jet clustering. At least two jets with $p_\mathrm{T} > 30\,\mathrm{GeV}/c$ and $|\eta| < 2.5$, separated by $\Delta R > 0.4$ from leptons passing the analysis selection, are required in each event. At least one of these jets is required to be consistent with coming from the decay of heavy flavor hadrons and be identified as a b jet by the Combined Secondary Vertex Medium Point (CSVM) b-tagging algorithm [9], which is based on the reconstruction of a secondary vertex. The $E_\mathrm{T}^\mathrm{miss}$ in the event is required to exceed 30 GeV, consistent with the presence of two undetected neutrinos.

Signal and background events are generated using the MADGRAPH 4.4.12 [10] and PYTHIA 6.4.22 [11] event generators, using next-to-leading order (NLO) cross sections. For $\mathrm{t\bar{t}}$ events, POWHEG with PYTHIA is used for the $\mathrm{t\bar{t}} \to \ell^+\ell^-$ component (corresponding to dileptonic $\mathrm{t\bar{t}}$, including τ leptons only when they also decay leptonically), while all other $\mathrm{t\bar{t}}$ decay modes, denoted $\mathrm{t\bar{t}} \to$ other, are generated using MADGRAPH. The samples of DY with $M_{\ell\ell} > 50\,\mathrm{GeV}/c^2$, diboson (WW, WZ, and ZZ only: the contribution from $\mathrm{W}\gamma$ is assumed to be negligible), and single top quark events are generated using MADGRAPH. The DY event samples with $M_{\ell\ell} < 50\,\mathrm{GeV}/c^2$ are generated using PYTHIA.

Events are then simulated using a GEANT4-based model [12] of the CMS detector, and finally reconstructed and analyzed with the same software used to process collision data.

With the steadily increasing LHC instantaneous luminosity, the mean number of interactions in a single bunch crossing also increased over the course of data taking, reaching about 15 at the end of the 2011 running period. In the following, the yields of simulated events are weighted such that the distribution of reconstructed vertices observed in data is reproduced. The efficiency for events containing two leptons satisfying the analysis selection to pass at least one of the double-lepton triggers is measured with a tag-and-probe method to be approximately 100%, 95%, and 90% for the ee, eμ, and $\mu\mu$ triggers, respectively [13], and corresponding weights are applied to the simulated event yields. In addition, b-tagging scale factors are applied to simulated events for each jet, to account for the difference between b-tagging efficiencies in data and simulation [9].

3 Event yields and top polarization at reconstruction level

The observed and simulated yields after the event selection are listed in Table 1. The yields are dominated (92%) by top-pair production in the dilepton final state, with the largest background coming from single top production. The $\mathrm{t\bar{t}} \to \ell^+\ell^-$ yields are normalized such that the total simulated yield matches the number of events in data. Comparisons between data and the simulation for the number of b-tagged jets and the number of vertices are shown in Figure 1.

From Equation 1, the top polarization can be extracted from

$$P_n = \frac{N(\cos(\theta_l^+) > 0) - N(\cos(\theta_l^+) < 0)}{N(\cos(\theta_l^+) > 0) + N(\cos(\theta_l^+) < 0)},$$

Table 1: The observed and simulated yields after the event selection described in the text. Uncertainties are statistical only. The systematic uncertainties on the simulated yields are given in Section 6. Where the simulated yields are zero, upper limits are given based on the weighted yield, had one of the simulated events passed the selection.

Sample	ee	$\mu\mu$	eμ	all
$t\bar{t} \to \ell^+\ell^-$	1791.7 ± 4.4	2127.3 ± 4.7	5069.4 ± 7.3	8988.5 ± 9.7
$t\bar{t} \to$ other	32.5 ± 2.9	4.8 ± 1.1	53.3 ± 3.6	90.7 ± 4.8
W + jets	< 1.9	4.7 ± 3.3	4.7 ± 3.4	9.4 ± 4.7
DY$\to$ ee	52.3 ± 5.8	< 0.6	< 0.6	52.3 ± 5.8
DY$\to \mu\mu$	< 0.6	72.8 ± 6.5	1.6 ± 0.9	74.4 ± 6.5
DY$\to \tau\tau$	17.6 ± 3.3	8.7 ± 2.2	18.7 ± 3.2	45.0 ± 5.1
Di-boson	10.6 ± 0.5	13.0 ± 0.5	24.0 ± 0.7	47.6 ± 1.0
Single top	84.9 ± 2.3	101.2 ± 2.4	252.1 ± 3.9	438.2 ± 5.1
Total (simulation)	1989.6 ± 8.8	2332.6 ± 9.3	5423.8 ± 10.3	9746.0 ± 16.4
Data	1961	2373	5412	9746

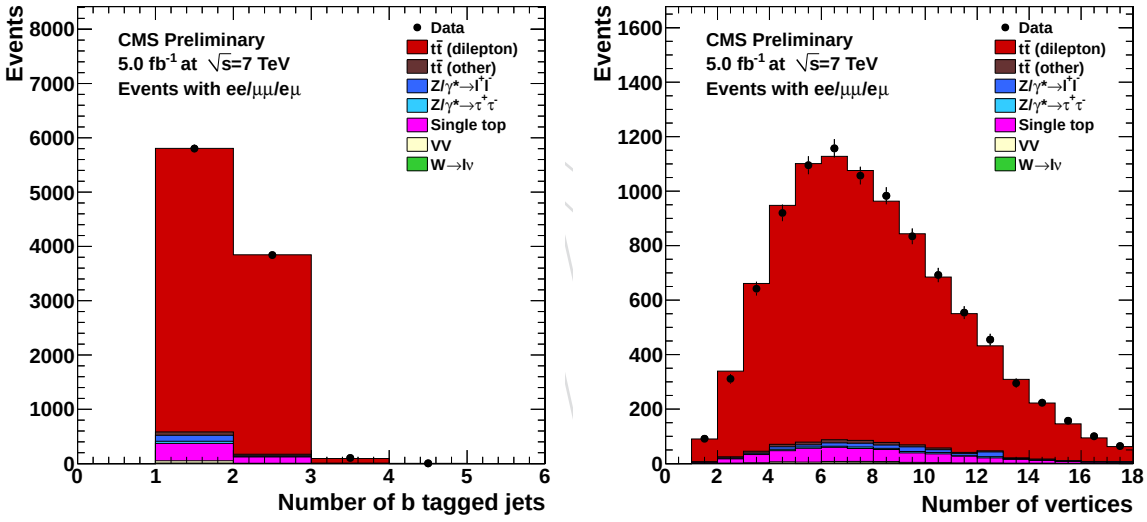

Figure 1: Comparison between data and the simulation for the number of b-tagged jets (left) and the number of vertices (right).

where θ_l^+ is the production angle of the positively charged lepton in the rest frame of its parent top, with respect to the direction of the parent top in the $t\bar{t}$ rest frame. Equivalently, P_n can be measured using θ_l^-, the analogous angle for the negatively charged lepton, but in this analysis θ_l^+ is used.

Measurement of this observable requires reconstruction of the $t\bar{t}$ system. Each event has two neutrinos, and there is also ambiguity in combining b jets with leptons. The analytical matrix weighting technique [14] is used to find a probable solution. Each event is reconstructed using a range of possible M_t values from 165-180 GeV/c^2 in 1 GeV/c^2 steps. The M_t hypothesis and the $t\bar{t}$ kinematics are then taken from the solution with largest weight. Approximately 33% of events have no solution found, and are thus not used in the measurement of the top polarization.

A comparison of the $\cos(\theta_l^+)$ distributions in data and the simulation is shown in Figure 2. The resulting value of P_n at reconstruction level is 0.040 ± 0.012 in data and 0.049 ± 0.002 in the simulation, where the uncertainties are statistical only. As a cross-check we also look at the $\cos(\theta_l^-)$ distributions at reconstruction level (also Figure 2), and find consistent results of 0.048 ± 0.012 in data and 0.046 ± 0.002 in the simulation. The remainder of this note focuses on making the corrections to the $\cos(\theta_l^+)$ distribution needed to obtain P_n at parton-level, necessary due to the presence of background events and resolution and acceptance effects.

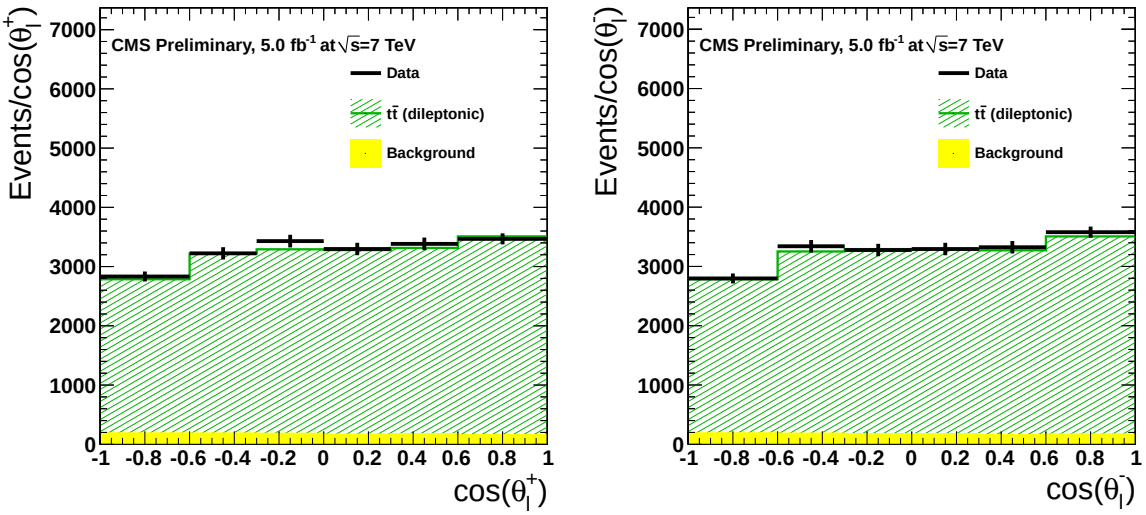

Figure 2: The reconstruction level $\cos(\theta_l^+)$ distribution (left) and $\cos(\theta_l^-)$ distribution (right), for data and the simulation. The simulated events are divided into $t\bar{t} \to \ell^+\ell^-$ and background, where the background consists of the categories other than $t\bar{t} \to \ell^+\ell^-$ in Table 1.

4 Background estimation

The simulation is used to predict the background event yields and shapes. We use methods based on data to cross-check these estimates for the background contributions from events with misidentified leptons and from DY$\to$ ee/$\mu\mu$ events. However, the dependence of the measured top polarization on the background normalization is small, and in Section 6 the systematic uncertainty is estimated based on changing the normalization of each component by up to 100%.

A misidentified lepton is defined as a lepton candidate not originating from a prompt decay, such as a lepton from semileptonic b or c quark decays, a muon from a pion or kaon decay, an unidentified photon conversion, or a pion misidentified as an electron. The background from events with misidentified leptons is predicted based on the number of events in data with a candidate lepton that can pass only loosened selection criteria [15]. Using a measurement of the fraction of such "loose" leptons that go on to pass the selection requirements, the number of misidentified leptons in the event sample can be estimated. The resulting prediction is 138^{+281}_{-138} events, including both statistical and systematic uncertainties. The simulated yield is 100.1 ± 6.7, in reasonable agreement.

The estimation method for DY$\to$ ee/$\mu\mu$ events [16] is based on counting the number of Z candidates in the Z veto region (after subtracting the number of non Z events estimated using the number of eμ events), and multiplying this number by the ratio of simulated DY yields

outside to inside the Z veto region. The result is an estimate of 142.4 ± 15.0 DY$\rightarrow$ ee/$\mu\mu$ events. The result is thus consistent with the simulated prediction of 126.7 ± 8.7 from Table 1.

5 Unfolding

The measured distribution is distorted from the true underlying distribution by the limited acceptance of our detector and the finite resolution of the measurement. In order to correct the data for these effects, we apply an unfolding procedure which yields the "parton-level" $\cos(\theta_l^+)$ distribution. This distribution represents the differential cross-section in $\cos(\theta_l^+)$, and is normalized to unity.

The choice of a binning scheme for the distribution is motivated by the desire to minimize bin-to-bin oscillations caused by statistical fluctuations. The bin size is variable and is chosen to ensure similar level of statistics in each bin of the distribution. A summary of the binning is provided in Table 2.

Table 2: Binning used in the distributions of $\cos(\theta_l^+)$.

B1	B2	B3	B4	B5	B6
[-1.0,-0.6]	[-0.6,-0.3]	[-0.3,-0.0]	[0.0, 0.3]	[0.3, 0.6]	[0.6, 1.0]

The background-subtracted measured distribution $\vec{b}$ is related to the underlying parton-level distribution $\vec{x}$ by the matrix equation $\vec{b} = SA\vec{x}$, where A is a diagonal matrix describing the acceptance in each bin of the measured distribution, and S is a non-diagonal smearing matrix describing the migration of events between bins due to the detector resolution and reconstruction techniques. The A and S matrices are modeled using the NLO POWHEG-PYTHIA $t\bar{t}$ sample, and are shown in Fig. 3. The smearing effects are quite large due to the uncertainties of top reconstruction. However, most of the large values lie close to the diagonal, meaning there is little migration between far-apart bins.

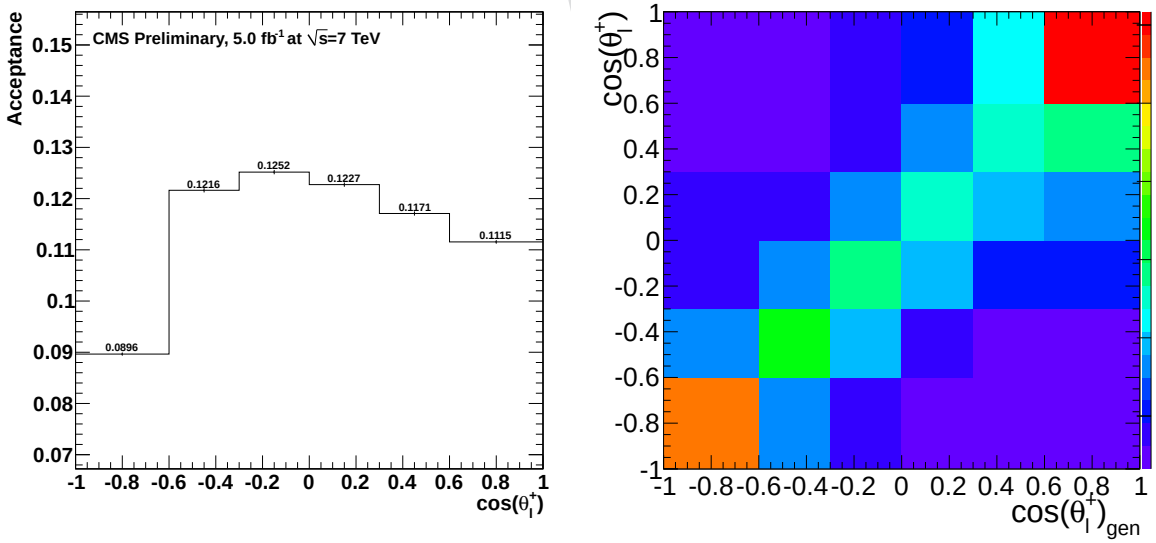

Figure 3: Acceptance matrix bins (left) and smearing effects due to the uncertainties of top reconstruction (right).

We employ a regularized "unfolding" algorithm based on singular value decomposition (SVD) [17], which is implemented in the RooUnfold package. The effects of large statistical fluctuations in the algorithm are greatly reduced by introducing a regularization term to the unfolding procedure. The full covariance matrix is used in the evaluation of the statistical uncertainty of the measured polarization.

We verify that the unfolding procedure is able to correctly unfold distributions with different levels of asymmetry. In order to do this, we re-weight generated $t\bar{t}$ events according to a linear function of $\cos(\theta_l^+)$: weight=1+$K\cos(\theta_l^+)$. The parameter K is varied between -0.5 and 0.5 in steps of 0.2, introducing a polarization of up to 40%, far more than is expected in $t\bar{t}$ events. For each value of K, we generate 2000 pseudo-experiments, in which the number of events in each bin of the distribution is fluctuated according to Poisson statistics, and then the distribution is unfolded. The average value of the asymmetry in 2000 pseudo-experiments is compared to the original true-level value. We find a linear behavior of this distribution, suggesting that non-SM asymmetry values will also be measured correctly. The offsets and slopes obtained in the linear function fit are -0.004 ± 0.009 and 1.031 ± 0.053, respectively. We also look at the distribution of the pulls in the set of pseudo-experiments and fit it to a Gaussian function. We find a small bias leading to asymmetry changes of up to 1%, the effect of which is included in the measurement uncertainty.

6 Systematic uncertainties

The systematic uncertainty associated with the jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ energy scale (JES) can directly affect the shape of the $\cos(\theta_l^+)$ distribution. We evaluate this uncertainty by varying the energy scale of jets within their uncertainty, which is parametrized as a function of the jet transverse energy and pseudorapidity, with the proper propagation to the $E_{\mathrm{T}}^{\mathrm{miss}}$ [18]. A similar uncertainty on the lepton energy scale is evaluated by shifting the electron energies by $\pm 0.3\%$ (the uncertainty on muon energies is negligible in comparison). We also evaluate a conservative systematic uncertainty on the choice of M_{t} scan range in the solver, taking the largest difference in result when changing the scan range from the default of 165-180 GeV/c^2 to ranges of 0-2500 GeV/c^2, 100-300 GeV/c^2, and a fixed value of 172.5 GeV/c^2.

The uncertainty associated with the background subtraction is obtained by changing the background yields. Given the uncertainties of the data driven background estimates described in Section 4, we vary the backgrounds from DY and misidentified leptons by 100%, and the single top background by 50%.

There are also several systematic uncertainties associated with the simulation of the events used to derive the unfolding matrices. The systematic uncertainty from $t\bar{t}$ modeling is estimated by applying unfolding derived using MC@NLO simulated $t\bar{t}$ events, and taking the difference from the default result using POWHEG-PYTHIA derived unfolding. Further systematic uncertainties are evaluated by applying unfolding derived using simulated events with the parameter of interest shifted up and down by an amount representing 1σ, and taking half the difference between the two results: the shower matching p_{T} thresholds are shifted by factors of 2 and 1/2; the factorization and renormalization scales (Q^2) are shifted from their default values of M_{t} down to $M_{\mathrm{t}}/2$ and up to $2M_{\mathrm{t}}$; the top quark mass used in the simulation is shifted down to 166.5 GeV/c^2 and up to 178 GeV/c^2; and the scale factors between data and simulation for $b-$tagging efficiency, trigger efficiency, and lepton ID efficiency are shifted up and down by 1σ. The systematic uncertainty from pile-up reweighting is verified to be small by applying unfolding derived using no pile-up reweighting.

The systematic uncertainties on the unfolded P_n measurement are summarized in Table 3, combining in quadrature to a total systematic uncertainty of 0.041.

Table 3: Systematic uncertainties.

JES	lepton energy scale	M_t scan range	background	$t\bar{t}$ modeling	matching
0.020	0.001	0.024	0.009	0.014	0.004
Q^2 scale	simulated M_t	$b-$tagging eff.	Trig eff. and lep ID	pile-up	Total
0.007	0.019	0.001	0.005	0.002	0.041

7 Results and summary

The background-subtracted and unfolded $\cos(\theta_l^+)$ distribution for $t\bar{t} \to \ell^+\ell^-$ events is shown in Figure 4, and is consistent with the parton-level prediction obtained from POWHEG-PYTHIA simulation. The measured value of P_n at parton-level is $-0.009 \pm 0.029 \pm 0.041$. The correlation between bins as a result of the unfolding procedure is accounted for in the evaluation of the uncertainties. The measured value of P_n is consistent with the POWHEG-PYTHIA value, and we thus observe no significant deviation from the SM expectation.

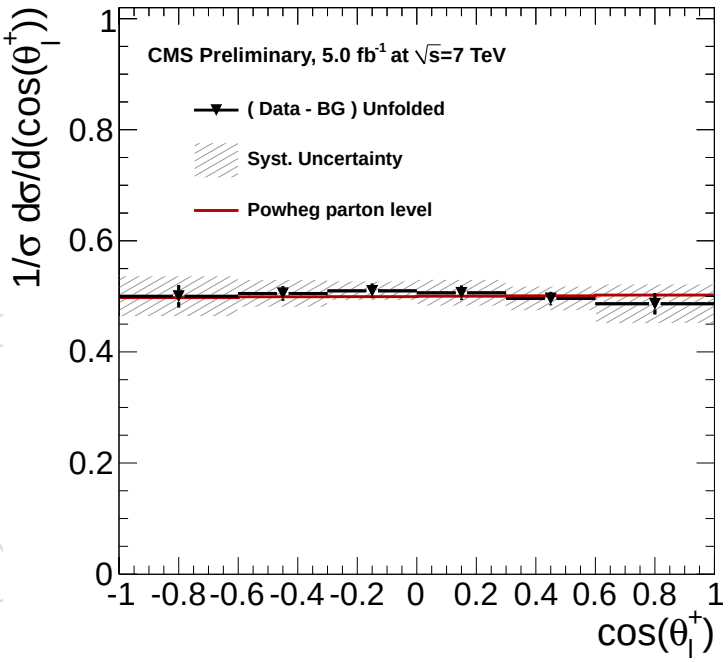

Figure 4: Background-subtracted and unfolded $\cos(\theta_l^+)$ distribution. The error bars represent statistical uncertainties only, while the systematic uncertainty band is represented by the hatched area. Note that the bin values are correlated due to the unfolding.

References

[1] D. Krohn, T. Liu, J. Shelton et al., "A Polarized View of the Top Asymmetry", *Phys.Rev.* **D84** (2011) 074034, doi:10.1103/PhysRevD.84.074034, arXiv:1105.3743.

[2] CDF Collaboration, "Forward-Backward Asymmetry in Top Quark Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV", *Phys.Rev.Lett.* **101** (2008) 202001, arXiv:0806.2472.

[3] D0 Collaboration, "First measurement of the forward-backward charge asymmetry in top quark pair production", *Phys.Rev.Lett.* **100** (2008) 142002, arXiv:0712.0851.

[4] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[5] CMS Collaboration, "Performance of muon identification in pp collisions at $\sqrt{s}$ = 7 TeV", CMS Physics Analysis Summary CMS-PAS-MUO-10-002, (2010).

[6] CMS Collaboration, "Electron Reconstruction and Identification at $\sqrt{s} = 7$ TeV", CMS Physics Analysis Summary CMS-PAS-EGM-10-004, (2010).

[7] CMS Collaboration, "Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV", CMS Physics Analysis Summary CMS-PAS-PFT-10-002, (2010).

[8] M. Cacciari, G. Salam, and G. Soyez, "The anti-k_T jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[9] CMS Collaboration, "Measurement of the b-tagging efficiency using $\mathrm{t\bar{t}}$ events", CMS Physics Analysis Summary CMS-PAS-BTV-11-003, (2011).

[10] J. Alwall, P. Demin, S. de Visscher et al., "MadGraph/MadEvent v4: the new web generation", *JHEP* **09** (2007) 028, doi:10.1088/1126-6708/2007/09/028, arXiv:0706.2334.

[11] T. Sjöstrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 physics and manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:0706.2334.

[12] S. Agostinelli et al., "GEANT4 – a simulation toolkit", *Nucl. Instr. and Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[13] CMS Collaboration, "Search for heavy, top-like quark pair production in the dilepton final state in pp collisions at $\sqrt{s} = 7$ TeV", *Submitted to PLB* (2012) arXiv:1203.5410.

[14] CMS Collaboration Collaboration, "Measurement of the t t-bar production cross section and the top quark mass in the dilepton channel in pp collisions at sqrt(s) =7 TeV", *JHEP* **1107** (2011) 049, arXiv:1105.5661.

[15] CMS Collaboration, "Search for new physics with same-sign isolated dilepton events with jets and missing transverse energy at the LHC", *JHEP* **6** (2011) 077, doi:10.1007/JHEP06(2011)077, arXiv:1104.3168.

[16] CMS Collaboration Collaboration, "First Measurement of the Cross Section for Top-Quark Pair Production in Proton-Proton Collisions at sqrt(s)=7 TeV", *Phys.Lett.* **B695** (2011) 424–443, arXiv:1010.5994.

[17] A. Hocker and V. Kartvelishvili, "SVD approach to data unfolding", *Nucl.Instrum.Meth.* **A372** (1996) 469–481, doi:10.1016/0168-9002(95)01478-0, arXiv:hep-ph/9509307.

[18] CMS Collaboration, "Determination of the Jet Energy Scale in CMS with pp Collisions at $\sqrt{s} = 7$ TeV", CMS Physics Analysis Summary CMS-PAS-JME-10-010, (2010).